(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 July 2004 (08.07.2004)

PCT

(10) International Publication Number
WO 2004/056964 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number:
PCT/US2003/040548

(22) International Filing Date:
18 December 2003 (18.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/434,856 18 December 2002 (18.12.2002) US
60/499,313 28 August 2003 (28.08.2003) US

(71) Applicant *(for all designated States except US)*: **GENPATH PHARMACEUTICALS, INCORPORATED** [US/US]; 300 Technology Square, 7th Floor, Cambridge, MA 02139 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **WINSTON, William** [US/US]; 108 Liberty Avenue, Apt. #1, Somerville, MA 02144 (US). **O'HAGAN, Ronan, C.** [CA/US]; 15 Garrison Road, Apt. 2, Brookline, MA 02445 (US). **HEYER, Joerg** [DE/US]; 6 Fainwood Circle, Cambridge, MA 02139 (US). **RIDEOUT, William** [US/US]; 20 Corporal Burns Road, Cambridge, MA 02142 (US). **ETEMAD-MOGHADAM, Bijan** [US/US]; 289 Vermont Street, West Roxbury, MA 02132 (US).

(74) Agents: **GUNNISON, Jane, T.** et al.; c/o Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: VECTORS FOR INDUCIBLE RNA INTERFERENCE

(57) Abstract: Recombinant vectors for inducibly expressing double-stranded RNA molecules that interfere with the expression of a target gene.

WO 2004/056964 A2

VECTORS FOR INDUCIBLE RNA INTERFERENCE

BACKGROUND OF THE INVENTION

RNA interference (RNAi) has been used to silence the expression of a target gene. RNAi is a sequence-specific posttranscriptional gene silencing mechanism triggered by double-stranded RNA (dsRNA). It causes degradation of mRNAs homologous in sequence to the dsRNA. The mediators of the degradation are 21- to 23-nucleotide small interfering RNAs (siRNAs) generated by cleavage of longer dsRNAs (including hairpin RNAs) by DICER, a ribonuclease III-like protein. Molecules of siRNA typically have 2- to 3-nucleotide 3' overhanging ends resembling the RNAse III processing products of long dsRNAs that normally initiate RNAi. When introduced into a cell, they assemble with an endonuclease complex (RNA-induced silencing complex), which then guides target mRNA cleavage. As a consequence of degradation of the targeted mRNA, cells with a specific phenotype of the suppression of the corresponding protein product are obtained (e.g., reduction of tumor size, metastasis, angiogenesis, and growth rates). The small size of siRNAs, compared with traditional antisense molecules, prevents activation of the dsRNA-inducible interferon system present in mammalian cells. This helps avoid the nonspecific phenotypes normally produced by dsRNA larger than 30 base pairs in somatic cells. See, e.g., Elbashir et al., Methods 26:199-213 (2002); McManus and Sharp, Nature Reviews 3:737-747 (2002); Hannon, Nature 418:244-251 (2002); Brummelkamp et al., Science 296:550-553 (2002); Tuschl,

Nature Biotechnology 20:446-448 (2002); U.S. Application US2002/0086356 A1; WO 99/32619; WO 01/36646; and WO 01/68836.

SUMMARY OF THE INVENTION

This invention features recombinant vectors for expressing double-stranded RNA (dsRNA) molecules in a controllable manner and cells and animals comprising the vectors. These dsRNA molecules interfere with (i.e., inhibit) the expression of a target gene, particularly a disease-related gene such as an oncogene or a tumor suppressor gene. In another aspect, the invention provides a modified nucleic acid molecule encoding a tetracycline repressor (TetR), for use in mammalian cells.

The vectors of this invention can be used to express the interfering dsRNAs at a desired time point to study the biological functions of a target gene *in vitro* or *in vivo*. Inducible expression of RNAi also has potential uses where RNAi is used as a therapeutic intervention to control the point of initiation and duration of expression of the therapeutic RNAi. The vectors can also be used to assess the toxicity of the dsRNAs *in vivo*, i.e., to determine whether expression of a given dsRNA has side effects in nontargeted cells and tissues and/or whether inhibition of a target gene causes undesired physiological problems.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Exemplary methods and materials are described below, although methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention. All publications and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. The materials, methods, and examples are illustrative only and not intended to be limiting. Throughout this specification, the word "comprise," or variations such as "comprises" or "comprising" will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

Other features and advantages of this invention will be apparent from the description below.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a drawing illustrating the construction of a vector system (constructs J-P, *infra*) that contains a tetracycline-inducible system for expressing an RNAi sequence. "TTTTT" denotes a stretch of nucleotide T that serves as a transcription terminator sequence.

Fig. 2A is a drawing illustrating several constructs (constructs A-E, *infra*) that contain an IPTG-inducible system for expressing an RNAi sequence. "TTTTT" denotes a stretch of nucleotide T that serves as a transcription terminator sequence.

Fig. 2B is a bar graph showing that modified human U6 promoters containing two Lac operator sequences (constructs C-E) are more strongly repressed by LacI than modified human U6 promoters containing only one Lac operator sequence (constructs A and B). "FF1" denotes an RNAi-encoding sequence that targets luciferase gene expression.

Fig. 2C is a bar graph showing that constructs A-E retain wildtype Lac operator function.

Fig. 3A is a diagram illustrating constructs F, G, H, and I (*infra*) of this invention.

Fig. 3B is a bar graph showing that removal of the LoxP-Stop-LoxP cassette in constructs F-I allows efficient transcription of the FF1 coding sequence.

Figs. 3C-E are drawings illustrating another construct that contains a Cre-inducible system for expressing an RNAi sequence. Upon activation of the Cre recombinase, the LoxP sites recombine, deleting the intervening STOPPER sequence.

Fig. 4A is a bar graph showing that constructs J-P (*infra*) of this invention retain wildtype Tet operator function.

Fig. 4B is a bar graph showing that modified human U6 promoters containing a TetO are repressed by TetR.

Fig. 5A is a bar graph showing that a modified human U6 promoter containing a TetO is repressed in a stable setting by a codon optimized TetR.

Fig. 5B is a line graph showing that RNAi of luciferase in xenograft tumors expressing (1) luciferase, (2) a luciferase shRNA from a modified human

U6 promoter TetO, and (3) a codon optimized TetR can be regulated by doxycycline.

Fig. 6 is a diagram illustrating a combined system of this invention, in which the LoxP system is combined with the TetO or LacO system.

Fig. 7 is a diagram illustrating the use of the Cre-Lox system to switch expression from one RNAi-encoding sequence (shRNA) to a second.

DETAILED DESCRIPTION OF THE INVENTION

This invention features recombinant vectors containing inducible systems for expressing dsRNA molecules that interfere with expression of target genes, including disease-related genes (e.g., cancer-related genes such as oncogenes and tumor suppressor genes). These vectors can be based on plasmids or viruses such as retroviruses (e.g., Moloney amphotropic murine virus), adenoviruses, and lentiviruses.

The vectors of this invention can be delivered into host cells via a variety of methods, including but not limited to, liposome fusion (transposomes), infection by viral vectors, and routine nucleic acid transfection methods such as electroporation, calcium phosphate precipitation and microinjection. Host cells include cultured cells and cells in an animal. In some embodiments, the vectors are integrated into the genome of a transgenic animal (e.g., a mouse, a rabbit, a hamster, or a nonhuman primate). Diseased or disease-prone cells containing these vectors can be used as a model system to study the development, maintenance, or progression of a disease that is affected by the presence or absence of the interfering RNA.

This model system can also be used to identify other disease-related elements. For instance, a detailed expression profile of gene expression in tumors undergoing regression or regrowth due to the expression or nonexpression of the interfering RNA can be established. Techniques used to establish the profile include the use of suppression subtraction (in cell culture), differential display, proteomic analysis, serial analysis of gene expression (SAGE), and expression/transcription profiling using cDNA and/or oligonucleotide microarrays. Then, comparisons of expression profiles at different stages of cancer development can be performed to identify genes whose expression patterns are altered.

Animals harboring the vectors of this invention can also be used to identify surrogate biomarkers for diagnosis or for following disease progression in patients. The biomarkers can be identified based on the differences between the expression profiles of the "on" and "off" states in the animal model. Blood or urine samples from the animal can be tested with ELISAs or other assays to determine which biomarkers are released from the diseased tissue (e.g., tumor) into circulation during genesis, maintenance, or regression of the disease. These biomarkers are particularly useful clinically in following disease progression post RNAi therapy or post-drug therapy which targets the same gene as the RNAi. These biomarkers can also be used clinically to assess the toxicity of any such therapy.

I. DESIGN OF VECTOR INSERTS

Useful interfering RNAs can be designed with a number of software programs, e.g., the OligoEngine siRNA design tool available at http://www.oligoengine.com. The siRNAs of this invention may range about, e.g., 19-29 basepairs in length for the double-stranded portion. In some embodiments, the siRNAs are hairpin RNAs having an about 19-29 bp stem and an about 4-34 nucleotide loop. Preferred siRNAs are highly specific for a region of the target gene and may comprise any about 19-29 bp fragment of a target gene mRNA that has at least one , preferably at least two or three, bp mismatch with a nontarget gene-related sequence. In some embodiments, the preferred siRNAs do not bind to RNAs having more than 3 mismatches with the target region.

Intracellular transcription of dsRNAs can be achieved by cloning the dsRNA-encoding sequences into RNA polymerase III (Pol III) transcription units, which normally encode the small nuclear RNA U6 or the human RNAse P RNA H1, or into RNA polymerase I (Pol I) or II (Pol II) transcription units (e.g., units containing a CMV promoter). However, it will be appreciated that in the vectors of the invention, the dsRNA-encoding sequences may be operably linked to a variety of other promoters. In some embodiments, the promoter is a convergent RNA polymerase III prompter, e.g., a convergent U6 snRNA promoter (Tran et al., *BMC Biotechnology* 3:21(2003)); a type II tRNA promoter such as the

$tRNA^{val}$ promoter and the $tRNA^{met}$ promoter. These promoters may also be modified to increase promoter activity. In addition, enhancers can be placed near the promoter to enhance promoter activity. For example, an enhancer from the CMV promoter can be placed near the U6 promoter to enhance U6 promoter activity (Xia et al., *Nuc Acids Res* 31 (2003)). Exemplary inducible Pol II systems are available from Invitrogen, e.g., the GeneSwitch™ and T-Rex™ systems.

Two approaches can be used for expressing dsRNA: (1) sense and antisense strands constituting the dsRNA duplex are transcribed by individual promoters; or (2) dsRNAs are expressed as fold-back stem-loop structures (hairpins) that give rise to dsRNAs after intracellular processing. Inducible transcription-regulatory elements are inserted into the promoter region for controlled expression of the dsRNAs. See, e.g., discussions below on the tetracycline-inducible, IPTG-inducible, and Cre-inducible Pol III-based transcription units. For Pol I- or Pol II-based transcription units, well-established inducible systems such as tetracycline transactivator systems, reverse tetracycline transactivator systems, and ecdysone systems can be used. However, it will be appreciated that for controlled expression of the dsRNAs, other operators that are controlled by a small molecule are useful in the vectors of the invention.

An exemplary human U6 transcription unit has the following sequence:

```
LOCUS       SP6-U6 genomic     860 bp    DNA
FEATURES             Location/Qualifiers
     snRNA           529..635
                     /label=U6 transcript
     enhancer        286..317
                     /label=Distal Sequence Element
     misc_binding    463..482
                     /label=Proximal Sequence Element
     TATA_signal     498..506
                     /label=TATA Box
     terminator      631..635
                     /label=Transcriptional Termination Signal
     promoter        1..18
                     /label=SP6 promoter
     misc_feature    51..528
                     /label=U6 promoter
     misc_feature    529..529
```

/label=Start of Transcription
misc_feature 51..860
/label=U6 genomic fragment
misc_feature 457..462
/label=NdeI

BASE COUNT 227 a 195 c 205 g 233 t
ORIGIN
1 atttaggtga cactatagaa tacaagcttg gctgcaggtc gacggatccc cccgagtcca
61 acacccgtgg gaatcccatg ggcaccatgg cccctcgctc caaaaatgct ttcgcgtcgc
121 gcagacactg ctcggtagtt tcggggatca gcgtttgagt aagagcccgc gtctgaaccc
181 tccgcgccgc cccggcccca gtggaaagac gcgcaggcaa aacgcaccac gtgacggagc
241 gtgaccgcgc gccgagcgcg cgccaaggtc gggcaggaag agggcctatt tcccatgatt
301 ccttcatatt tgcatatacg atacaaggct gttagagaga taattagaat taatttgact
361 gtaaacacaa agatattagt acaaaatacg tgacgtagaa agtaataatt tcttgggtag
421 tttgcagttt taaaattatg ttttaaaatg gactatcata tgcttaccgt aacttgaaag
481 tatttcgatt tcttggctttt atatatcttg tggaaaggac gaaacaccgt gctcgcttcg
541 gcagcacata tactaaaatt ggaacgatac agagaagatt agcatggccc ctgcgcaagg
601 atgacacgca attcgtgaa gcgttccata tttttacatc aggttgtttt tctgttttta
661 catcaggttg tttttctgtt tggtttttt tttacaccac gtttatacgc cggtgcacgg
721 tttaccactg aaaacacctt tcatctacag gtgatatctt ttaacacaaa taaaatgtag
781 tagtcctagg agacggaata gaaggaggtg gggcctaggc agattcatct ctgcggtgca
841 ttttgcctct ggccctcggg (SEQ ID NO:1)

In this sequence, the U6-transcript sequence is underlined, with the PSE italicized and the transcription initiation site G double-underlined. The promoter region spans from nucleotide 51 to the nucleotide immediately preceding the initiating G. The TATA box in the promoter region is boxed. Luukkonen et al., RNA 4:231-8

(1998). To construct an RNAi vector, the U6-transcript sequence can be replaced in part or in its entirety by a sequence encoding an interfering dsRNA. In some embodiments, it may be preferred that the spacing between the PSE and the TATA box and the spacing between the TATA box and the GTG site are maintained for proper transcription. The cytosine (C) immediately preceding the GTG site and the purine immediately following this site may also be preserved for proper start of transcription. Goomer et al., Nucleic Acids Research 20:4903-12 (1992).

II. A TETRACYCYLINE-INDUCIBLE SYSTEM

Fig. 1 illustrates one vector system of this invention (see Working Examples). To construct this vector, a Tet operator sequence (TetOp) is inserted into the promoter region of the vector. TetOp is preferably inserted between the PSE and the transcription initiation site, upstream or downstream from the TATA box. In some embodiments, the TetOp is immediately adjacent to the TATA box.

The expression of the RNAi molecule is thus under the control of tetracycline (or doxycycline, or any other tetracycline analogue). Addition of tetracycline relieves repression of the promoter by a tetracycline repressor that the host cells are also engineered to express. Since the tetracycline repressor is derived from bacteria, its coding sequence may be optionally modified to adapt to the codon usage by mammalian transcriptional systems and to prevent methylation. In some embodiments, the host cells comprise (i) a first expression construct containing a gene encoding a tetracycline repressor operably linked to a first promoter, such as any tissue or cell type-specific promoter or any general promoter, and (ii) a second expression construct containing the dsRNA-coding sequence operably linked to a second promoter that is regulated by the tetracycline repressor and tetracycline. Administration of tetracycline or an analogue thereof (e.g., doxycycline) results in expression of the dsRNA in a manner dictated by the tissue specificity of the first promoter.

III. A LAC OPERATOR SYSTEM

Fig. 2A illustrates yet another vector system of this invention (see Working Examples). To construct this vector, a Lac operator sequence (LacO) is inserted into the promoter region. The LacO is preferably inserted between the

PSE and the transcription initiation site, upstream or downstream of the TATA box. In some embodiments, the LacO is immediately adjacent to the TATA box.

The expression of the RNAi molecule is thus under the control of IPTG (or any analogue thereof). Addition of IPTG relieves repression of the promoter by a Lac repressor (i.e., the LacI protein) that the host cells are also engineered to express. Since the Lac repressor is derived from bacteria, its coding sequence may be optionally modified to adapt to the codon usage by mammalian transcriptional systems and to prevent methylation. In some embodiments, the host cells comprise (i) a first expression construct containing a gene encoding a Lac repressor operably linked to a first promoter, such as any tissue or cell type-specific promoter or any general promoter, and (ii) a second expression construct containing the dsRNA-coding sequence operably linked to a second promoter that is regulated by the Lac repressor and IPTG. Administration of IPTG results in expression of dsRNA in a manner dictated by the tissue specificity of the first promoter.

IV. A LOXP-STOP-LOXP SYSTEM

Figs. 3A-E illustrate yet another vector system of this invention. The RNAi vector of this system contains a LoxP-Stop-LoxP cassette before the hairpin (Fig. 3A) or within the loop of the hairpin (Figs. 3C-E). Any suitable stop sequence for the promoter can be used in the cassette. One version of the LoxP-Stop-LoxP system for Pol II is described in, e.g., Wagner et al., Nucleic Acids Research 25:4323-4330 (1997). The "Stop" sequences (such as the one described in Wagner, *supra*, or a run of five or more T nucleotides) in the cassette prevent the RNA polymerase III from extending an RNA transcript beyond the cassette.

Upon introduction of a Cre recombinase, however, the LoxP sites in the cassette recombine, removing the Stop sequences and leaving a single LoxP site. Removal of the Stop sequences allows transcription to proceed through the hairpin sequence, producing a transcript that can be efficiently processed into an open-ended, interfering dsRNA. Thus, expression of the RNAi molecule is induced by addition of Cre.

In some embodiments, the host cells contain a Cre-encoding transgene under the control of a constitutive, tissue-specific promoter. As a result, the interfering RNA can only be inducibly expressed in a tissue-specific manner dictated by that promoter. Tissue-specific promoters that can be used include, without limitation: a tyrosinase promoter or a TRP2 promoter in the case of melanoma cells and melanocytes; an MMTV or WAP promoter in the case of breast cells and/or cancers; a Villin or FABP promoter in the case of intestinal cells and/or cancers; a RIP promoter in the case of pancreatic beta cells; a Keratin promoter in the case of keratinocytes; a Probasin promoter in the case of prostatic epithelium; a Nestin or GFAP promoter in the case of CNS cells and/or cancers; a Tyrosine Hydroxylase, S100 promoter or neurofilament promoter in the case of neurons; the pancreas-specific promoter described in Edlund et al. Science 230:912-916 (1985); a Clara cell secretory protein promoter in the case of lung cancer; and an Alpha myosin promoter in the case of cardiac cells.

Cre expression also can be controlled in a temporal manner, e.g., by using an inducible promoter, or a promoter that is temporally restricted during development such as Pax3 or Protein O (neural crest), Hoxa1 (floorplate and notochord), Hoxb6 (extraembryonic mesoderm, lateral plate and limb mesoderm and midbrain-hindbrain junction), Nestin (neuronal lineage), GFAP (astrocyte lineage), Lck (immature thymocytes). Temporal control also can be achieved by using an inducible form of Cre. For example, one can use a small molecule controllable Cre fusion, for example a fusion of the Cre protein and the estrogen receptor (ER) or with the progesterone receptor (PR). Tamoxifen or RU486 allow the Cre-ER or Cre-PR fusion, respectively, to enter the nucleus and recombine the LoxP sites, removing the LoxP Stop cassette. Mutated versions of either receptor may also be used. For example, a mutant Cre-PR fusion protein may bind RU486 but not progesterone. Other exemplary Cre fusions are a fusion of the Cre protein and the glucocorticoid receptor (GR). Natural GR ligands inlcude corticosterone, cortisol, and aldosterone. Mutant versions of the GR receptor, which respond to, e.g., dexamethasone, triamcinolone acetonide, and/or RU38486, mya also be fused to the Cre protein.

V. CELLS AND ANIMALS

This invention also provides nonhuman transgenic animals whose somatic and germ cells contain an inducible RNAi contsruct of this invention (including both heterozygotes and homozygotes). Such animals can be used to study the effect of the RNAi coding sequence on tumorigenicity and tumor development, to study the role of the targeted gene in normal tissue development and differentiation, and to screen for and establish toxicity profiles of anti-cancer drugs. Also included are chimeric animals that can be used to generate the transgenic animals. The non-human animal is preferably a mammal, more preferably a cow, goat, sheep, or rodent such as a rat or mouse.. As used herein, a "chimeric animal" is one in which one or more of the cells of the animal includes a transgene. In other embodiments, the transgenic or chimeric animals can be non-human primates, dogs, chickens, amphibians, etc.

VI. WORKING EXAMPLES

The following examples are meant to illustrate the methods and materials of the present invention. Suitable modifications and adaptations of the described conditions and parameters normally encountered in the art which are obvious to those skilled in the art are within the spirit and scope of the present invention.

A. IPTG-Inducible U6 Promoters

The following describes several IPTG-inducible U6 promoters that were made. All of these constructs encoded a small interfering RNA molecule that inhibited expression of a luciferase gene. Four versions of the *E. coli* LacO sequence were used. Short and long versions of the natural LacO sequence were used, 5'-aattgtgagcggataacaatt-3(SEQ ID NO:2) and 5'-tgtgtggaattgtgagcggataacaatttcacaca-3' (SEQ ID NO:3), respectively. Also, short and long versions of a synthetic LacO sequence (a perfect palindrome of the 5' half of the natural LacO) were used, 5'-gaattgtgagcgctcacaattc-3' (SEQ ID NO:4) and 5'-tgtggaattgtgagcgctcacaattccaca-3' (SEQ ID NO:5), respectively. In construct A (also called U6 LOM FF1), a short version of the natural *E. coli* LacO sequence (SEQ ID NO:2; *supra*) was inserted between the TATA box and the transcription

initiation site, replacing a native sequence of the same length (Fig. 2A). Construct A has the following sequence in the promoter region and the siRNA-coding region:

```
LOCUS       U6 LOM FF1    636 bp    DNA
FEATURES             Location/Qualifiers
     precursor_RNA   563..625
                     /label=FF1 shRNA
     misc_feature    536..562
                     /label=U6 leader sequence
     terminator      626..630
                     /label=Termination Signal
     misc_feature    58..535
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    536..536
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    470..489
                     /label=Proximal Sequence Element
     tRNA            505..513
                     /label=TATA Box
     misc_binding    514..534
                     /label=Natural Lac Operator
     misc_feature    1..6
                     /label=XhoI
     misc_feature    631..636
                     /label=EcoRI
     misc_feature    464..469
                     /label=NdeI
BASE COUNT      166 a      150 c      158 g      162 t
ORIGIN
        1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac
ggatcccccc
       61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa
aaatgctttc
      121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag
agcccgcgtc
      181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac
gcaccacgtg
      241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg
gcctatttcc
```

301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa

361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct

421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatatgc *ttaccgtaac*

481 *ttgaaagtat* ttcgatttct tggc**tttata tat**AATTGTG AGCGGATAAC AATTcgtgct

541 cgcttcggca gcacatatac taggattcca attcagcggg agccacctga tttggatcgg

601 gtggctctcg ctgagttgga atccattttt gaattc (SEQ ID NO:6)

In this sequence, the PSE is italicized; the TATA box is boxed; the LacO sequence is in uppercase; and the G transcription initiation site is double-underlined.

In construct B (also U6 LOP FF1), a short synthetic LacO sequence (SEQ ID NO: 4, *supra*) was inserted into the U6 promoter region in the same fashion. Construct B has the following sequence:

```
LOCUS       U6 LOP FF1     636 bp    DNA
FEATURES             Location/Qualifiers
     precursor_RNA   563..625
                      /label=FF1 shRNA
     misc_feature    536..562
                     /label=U6 leader sequence
     terminator      626..630
                     /label=Termination Signal
     misc_feature    58..513
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    536..536
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    470..489
                     /label=Proximal Sequence Element
     tRNA            505..513
                     /label=TATA Box
     misc_binding    514..535
                     /label=Synthetic Lac Operator
     misc_feature    1..6
                     /label=XhoI
     misc_feature    631..636
                     /label=EcoRI
```

misc_feature 464..469
/label=NdeI
BASE COUNT 164 a 152 c 158 g 162 t
ORIGIN
1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc
61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc
121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc
181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg
241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc
301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa
361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct
421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatatgc *ttaccgtaac*
481 *ttgaaagtat* ttcgatttct tggctttata tatGAATTGT GAGCGCTCAC AATTCgtgct
541 cgcttcggca gcacatatac taggattcca attcagcggg agccacctga tttggatcgg
601 gtggctctcg ctgagttgga atccattttt gaattc (SEQ ID NO:7)

In the above sequence, the markings are the same as in construct A's sequence.

Constructs A and B both showed wild-type transcriptional activity when LacI was not present, and showed repressed transcription activity when LacI was present.

To take advantage of LacI's ability to cooperatively bind two LacO sites, three more constructs, constructs C-E, were made, each containing two LacO sites. Construct C (also called U6 LO NdeI S LOM FF1) had the long synthetic LacO sequence inserted at the NdeI site and the short natural LacO sequence inserted between the TATA box and the G initiation site. It has the following sequence:

LOCUS LO NdeI S LOM FF1 670 bp DNA
FEATURES Location/Qualifiers
precursor_RNA 597..659
/label=FF1 shRNA
misc_feature 570..596

```
                    /label=U6 leader sequence
     terminator     660..664
                    /label=Termination Signal
     misc_feature   58..569
                    /label=U6 Promoter
     promoter       8..25
                    /label=SP6 Promoter
     misc_feature   570..570
                    /label=Start of Transription
     enhancer       293..324
                    /label=Distal Sequence Element
     misc_feature   504..523
                    /label=Proximal Sequence Element
     tRNA           539..547
                    /label=TATA Box
     misc_binding   548..568
                    /label=Natural Lac Operator
     misc_binding   469..498
                    /label=Synthetic Lac Operator (LO NdeI S)
     insertion_seq  466..499
                    /label=LO NdeI S insert
     misc_feature   1..6
                    /label=XhoI
     misc_feature   665..670
                    /label=EcoRI
BASE COUNT      176 a       157 c       165 g       172 t
ORIGIN
        1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac
ggatcccccc
       61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa
aaatgctttc
      121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag
agcccgcgtc
      181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac
gcaccacgtg
      241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg
gcctatttcc
      301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa
ttagaattaa
      361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt
aataatttct
      421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatattg
tggaattgtg
```

481 agcgctcaca attccacaat atgcttaccg taacttgaaa gtatttcgat ttcttggctt

541 tatatatAAT TGTGAGCGGA TAACAATTcg tgctcgcttc ggcagcacat atactaggat

601 tccaattcag cgggagccac ctgatttgga tcgggtggct ctcgctgagt tggaatccat

661 ttttgaattc (SEQ ID NO:8)

In the above sequence, the long synthetic LacO sequence is in boldface; and the short natural LacO sequence is in uppercase.

In construct D (also called U6 LO NdeI N LOM FF1), the long natural LacO was inserted at the NdeI site, and the short natural LacO was inserted between the TATA box and the G initiation site. This construct has the following sequence:

```
LOCUS       LO NdeI N LOM FF1    675 bp    DNA
FEATURES             Location/Qualifiers
     precursor_RNA   602..664
                     /label=FF1 shRNA
     misc_feature    575..601
                     /label=U6 leader sequence
     terminator      665..669
                     /label=Termination Signal
     misc_feature    58..574
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    575..575
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    509..528
                     /label=Proximal Sequence Element
     tRNA            544..552
                     /label=TATA Box
     misc_binding    553..573
                     /label=Natural Lac Operator
     insertion_seq   466..504
                     /label=LO NdeI N insert
     misc_binding    469..503
                     /label=Natural Lac Operator (LO NdeI N)
     misc_feature    670..675
                     /label=EcoRI
```

misc_feature 1..6
/label=XhoI
BASE COUNT 179 a 155 c 167 g 174 t
ORIGIN
1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc
61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc
121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc
181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg
241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc
301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa
361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct
421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatattg tgtggaattg
481 tgagcggata acaatttcac acaatatgct *taccgtaact* tgaaagtatt tcgatttctt
541 ggctttatat atAATTGTGA GCGGATAACA ATTcgtgctc gcttcggcag cacatatact
601 aggattccaa ttcagcggga gccacctgat ttggatcggg tggctctcgc tgagttggaa
661 tccatttttg aattc (SEQ ID NO:9)

In the above sequence, the long natural LacO sequences is in boldface, and the short natural LacO sequence is in uppercase.

In construct E (also called U6 LO NdeI S LOP FF1), the long synthetic LacO was inserted at the NdeI site, and the short synthetic LacO was inserted between the TATA box and the G initiation site. This construct has the following sequence:

LOCUS LO NdeI S LOP 670 bp DNA
FEATURES Location/Qualifiers
precursor_RNA 597..659
/label=FF1 shRNA
misc_feature 570..596
/label=U6 leader sequence
terminator 660..664
/label=Termination Signal
misc_feature 58..547

/label=U6 Promoter

promoter 8..25

/label=SP6 Promoter

misc_feature 570..570

/label=Start of Transription

enhancer 293..324

/label=Distal Sequence Element

misc_feature 504..523

/label=Proximal Sequence Element

tRNA 539..547

/label=TATA Box

misc_binding 548..569

/label=Synthetic Lac Operator

misc_binding 469..498

/label=Synthetic Lac Operator (LO NdeI S)

insertion_seq 466..499

/label=LO NdeI S insert

misc_feature 1..6

/label=XhoI

misc_feature 665..670

/label=EcoRI

BASE COUNT 174 a 159 c 165 g 172 t

ORIGIN

1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc

61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc

121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc

181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg

241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc

301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa

361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct

421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatattg tggaattgtg

481 agcgctcaca attccacaat atgcttaccg taacttgaaa gtatttcgat ttcttggctt

541 tatatatGAA TTGTGAGCGC TCACAATTCg tgctcgcttc ggcagcacat atactaggat

601 tccaattcag cgggagccac ctgatttgga tcgggtggct ctcgctgagt tggaatccat

661 ttttgaattc (SEQ ID NO:10)

In the above sequences, the long synthetic LacO sequence is in boldface, and the short synthetic LacO sequence is in uppercase.

In all of constructs A-E, the spacing was not changed between the TATA box and the G initiation site when the LacO sequence was inserted therein. That is, the LacO sequence replaced a native sequence of the same length. The following shows an alignment of the promoter sequences of wildtype U6 and constructs A-E. The numbers above the sequences denote nucleotide positions, where position "1" corresponds to nucleotide 444 in SEQ ID NO:1. The following sequences are assigned SEQ ID NOs:11-16, respectively.

```
          1                                                                65
U6 WT     taaaatggactatca----------------------------------------tatgcttaccg
A         taaaatggactatca----------------------------------------tatgcttaccg
B         taaaatggactatca----------------------------------------tatgcttaccg
C         taaaatggactatcatat--tgtggaattgtgagc-gctcacaattccaca--atatgcttaccg
D         taaaatggactatcatattgtgtggaattgtgagcggataacaatttcacacaatatgcttaccg
E         taaaatggactatcatat--tgtggaattgtgagc-gctcacaattccaca--atatgcttaccg
          66                                                              127
U6 WT     taacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggacgaaacaccg
A         taacttgaaagtatttcgatttcttggctttatatatAATTGTGAGCGGATAACAATTcg
B         taacttgaaagtatttcgatttcttggctttatatatGAATTGTGAGCGCTCACAATTCg
C         taacttgaaagtatttcgatttcttggctttatatatAATTGTGAGCGGATAACAATTcg
D         taacttgaaagtatttcgatttcttggctttatatatAATTGTGAGCGGATAACAATTcg
E         taacttgaaagtatttcgatttcttggctttatatatGAATTGTGAGCGCTCACAATTCg
```

Also, in all of constructs A-E, the transcription termination site is "TTTTT" following the G initiation site and the short-hairpin sequence. The siRNA coding sequence encodes a short-hairpin sequence (FF1) targeting Luciferase expressed from the pGL3 vector (Promega).

Constructs A-E were constructed as described below. For construct A, to insert the LacO sequence into the promoter region, the following primer sets were used in two separate polymerase chain reactions (PCR): (1) a first external primer (5'-ggccctcgaggatttaggtgacactatag-3'; SEQ ID NO:17) that targeted a vector region 5' to the U6 transcription unit, and a first internal primer having the sequence of 5'-agcacgaattgttatccgctcacaattatatataaagccaagaaatcgaaatact-3' (SEQ

ID NO:18); and (2) a second internal primer having the sequence of 5'-tggctttatatataattgtgagcggataacaattcgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:19), and a second external primer that targeted a vector region 3' to the U6 transcription unit. The PCR products from these two reactions were mixed and subjected to PCR again using the two external primers. This PCR reaction generated a complete, modified U6 transcription unit having the LacO sequence. To construct an RNAi vector, the U6-coding sequence was then replaced in part or in its entirety by a sequence encoding an interfering dsRNA, using routine recombinant techniques (e.g., PCR).

For construct B, to insert the LacO sequence into the promoter region, the following primer sets were used in two PCRs: (1) a first external primer (SEQ ID NO:17, *supra*) that targeted a vector region 5' to the U6 transcription unit, and a first internal primer having the sequence of 5'-agcacgaattgtgagcgctcacaattcatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:20); and (2) a second internal primer having the sequence of 5'-tggctttatatatgaattgtgagcgctcacaattcgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:21), and a second external primer that targeted a vector region 3' to the U6 transcription unit. The PCR products from these two reactions were mixed and subjected to PCR again using the two external primers. This PCR reaction generated a complete, modified U6 transcription unit having the LacO sequence. To construct an RNAi vector, the U6-coding sequence was replaced in part or in its entirety by a sequence encoding an interfering dsRNA, using routine recombinant techniques (e.g., PCR).

For constructs A and B, the FF1 sequence with an EcoRI site was introduced into the human U6 promoter DNA sequence using PCR with an XhoI-containing forward primer (SEQ ID NO:17, *supra*) and the FF1-containing reverse primer 5'-ggaattcaaaaatggattccaactcagcgagagccacccgatccaaatcaggtggctcccgctgaattggaatcctagtatatgtgctgccgaagc-3' (SEQ ID NO:22). U6 promoter DNA fragments containing the FF1 sequence were digested with EcoRI and XhoI and inserted into pENTR11 (Invitrogen) digested with EcoRI and XhoI.

Construct C was made by digesting construct A with NdeI and inserting the annealed self-complementary oligonucleotide 5'-tattgtggaattgtgagcgctcacaattccacaa-3' (SEQ ID NO:23), introducing the long synthetic LacO sequence.

Construct D was made by digesting construct A with NdeI and inserting the annealed oligonucleotides 5'-tattgtgtggaattgtgagcggataacaatttcacacaa-3' (SEQ ID NO:24) and 5'-tattgtgtgaaattgttatccgctcacaattccagaca-3' (SEQ ID NO:25), introducing the long natural LacO sequence.

Construct E was made by digesting construct B with NdeI and inserting the annealed self-complementary oligonucleotide SEQ ID NO:23 (*supra*), introducing the long synthetic LacO sequence.

The constructs were then cotransfected into NIH 3T3 cells with pGL3 Control (for directing luciferase expression; Promega), pCMV LacI (for direction LacI expression; Stratagene), and pSEAP2 Control (for directing SEAP expression; as a control for co-transfection; BD Biosciences). Transfected cells were treated with 5 mM IPTG and were compared to untreated cells.

Forty-eight hours after the start of transfection, 10 μl of supernatant from the cells was removed and used for a SEAP luminescence assay (Great ESCAPE SEAP Chemiluminescence Assay, BD Biosciences). The cells were subjected to a luciferase luminescence assay (STEADY-GLO, Promega). Data were normalized for transfection efficiency by dividing luciferase assay values with SEAP assay values. The comparison between transfection with the empty vector (pENTR11) and transfection with U6 FF1 vectors shows the degree of inhibition of luciferase expression from pGL3.

As shown in Fig. 2B, U6 constructs containing two LacO sequences (i.e., constructs C-E), when co-transfected with pCMV LacI, exhibited stronger (approximately 15% more) repression by LacI in the absence of IPTG (i.e., thus less expression of the FF1 transcript and less inhibition of luciferase expression), compared to U6 constructs containing only one LacO sequence (e.g., constructs A-B). Fig. 2C shows that constructs A-E inhibited luciferase expression significantly in the absence of LacI expression. The extents of the inhibition among the constructs were comparable. Constructs A and B had a combined average

inhibition of 91%, and constructs C-E had a combined average inhibition of 83%. In conclusion, U6 constructs containing two LacO sequences retained wild type promoter activity, while having tighter inducibility control by LacI.

B. Cre-LoxP Systems

The Cre-LoxP system was used to create U6 promoter constructs for tissue-specific expression of short-hairpin RNAs (shRNAs) that target the expression of a luciferase gene. A LoxP-Stop-LoxP cassette was inserted between the G transcriptional start site of the U6 promoter and the FF1 shRNA-coding sequence (Fig. 3A). The approximately 400 base pair LoxP Stop cassette consisted of two LoxP sites in the same orientation bracketing six RNA Polymerase III transcriptional termination sites (stretches of four or more Ts from a luciferase gene fragment and a U6 RNA transcriptional termination fragment). The LoxP Stop cassette prevented transcription from proceeding through to the shRNA-coding sequence. Cre-mediated recombination would remove the intervening Stop sequence between the two LoxP sites, leaving only one LoxP site. Transcription could then proceed through to the shRNA-coding sequence.

Four U6 LoxP Stop constructs were made, termed constructs F, G, H, and I. The differences among these constructs were the position of the transcriptional start site relative to the first LoxP site and the sequence used as a spacer between the second LoxP site and the FF1 sequence. In construct F (also U6 LoxP Stop 1A FF1), the initiation start site was placed directly 5' to the first LoxP site, while the spacer between the second LoxP site and the FF1 sequence was 5'-CGACGAGGC-3'. Construct G (also U6 LoxP Stop 1B FF1) was identical to construct F except that it had a spacer sequence of 5'-CGACCTCCC-3'. In construct H (also U6 LoxP Stop 2A FF1), the transcription initiation start site was placed within the first LoxP site, maintaining the U6 transcriptional nucleotide as G, preceded by C as in the wild-type promoter. This construct had the same spacer sequence as construct F. Construct I (also U6 LoxP Stop 2B FF1) was the same as construct H except that the former had the spacer sequence of construct G.

In order to compare the U6 LoxP constructs to their recombined form (i.e., with the STOPPER sequence removed), constructs F, G, H, and I were each recombined *in vitro* using the Cre recombinase. The ability of the cloned,

recombined construct to silence Luciferase was then assayed and compared to a wild-type U6 promoter expressing the FF1 shRNA targeting Luciferase.

The promoter sequence of construct F is shown below.

```
LOCUS       U6 Lox Stop Lox 1A FF1   1021 bp    DNA
FEATURES             Location/Qualifiers
     misc_feature    58..535
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    536..536
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    470..489
                     /label=Proximal Sequence Element
     tRNA            505..513
                     /label=TATA Box
     misc_recomb     537..570
                     /label=LoxP
     misc_recomb     905..938
                     /label=LoxP
     terminator      620..625
                     /label=Transcriptional Termination 1
     misc_feature    571..739
                     /label=pGL3 Luciferase fragment
     terminator      753..904
                     /label=U6 Transcriptional Termination
     precursor_RNA   948..1010
                     /label=FF1 shRNA
     terminator      1011..1015
                     /label=Transcriptional Termination 7
     misc_feature    1016..1021
                     /label=EcoRI
     misc_feature    1..6
                     /label=XhoI
     terminator      789..793
                     /label=Transcriptional Termination 2
     terminator      805..809
                     /label=Transcriptional Termination 3
     misc_feature    813..817
                     /label=Transcriptional Termination 4
     terminator      829..833
                     /label=Transcriptional Termination 5
```

terminator 842..851
/label=Transcriptional Termination 6
BASE COUNT 265 a 223 c 253 g 280 t
ORIGIN
1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc
61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc
121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc
181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg
241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc
301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa
361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct
421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatatgc *ttaccgtaac*
481 *ttgaaagtat* ttcgatttct tggctttata tatcttgtgg aaaggacgaa acaccgataa
541 **cttcgtatag catacattat** acgaagttat tacacccgag ggggatgata aaccgggcgc
601 ggtcggtaaa gttgttccat tttttgaagc gaaggttgtg gatctggata ccgggaaaac
661 gctgggcgtt aatcaaagag gcgaactgtg tgtgagaggt cctatgatta tgtccggtta
721 tgtaaacaat ccggaagcgc cgcggccgct aggcaaggat gacacgcaaa ttcgtgaagc
781 gttccatatt tttacatcag gttgtttttc tgtttttaca tcaggttgtt tttctgtttg
841 gtttttttt tacaccacgt ttatacgccg gtgcacggtt taccactgaa aacacctttc
901 atctataact **tcgtatagca tacattatac gaagttatcg** acgaggcgga ttccaattca
961 gcgggagcca cctgatttgg atcgggtggc tctcgctgag ttggaatcca tttttgaatt
1021 c (SEQ ID NO:26)

In the above sequence, the PSE is italicized; the LoxP sites are in boldface; and the initiation start site is double-underlined. After recombination, only one LoxP site

remains and the sequence is the same as if the G nucleotide at position 536 is joined immediately 5' to the A nucleotide at position 905.

The promoter sequence of construct G is shown below.

```
LOCUS       U6 Lox Stop Lox 1B FF1   1021 bp    DNA
FEATURES             Location/Qualifiers
     misc_feature    58..535
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    536..536
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    470..489
                     /label=Proximal Sequence Element
     tRNA            505..513
                     /label=TATA Box
     misc_recomb     537..570
                     /label=LoxP
     misc_recomb     905..938
                     /label=LoxP
     terminator      620..625
                     /label=Transcriptional Termination 1
     misc_feature    571..739
                     /label=pGL3 Luciferase fragment
     terminator      753..904
                     /label=U6 Transcriptional Termination
     precursor_RNA   948..1010
                     /label=FF1 shRNA
     terminator      1011..1015
                     /label=Transcriptional Termination 7
     misc_feature    1016..1021
                     /label=EcoRI
     misc_feature    1..6
                     /label=XhoI
     terminator      789..793
                     /label=Transcriptional Termination 2
     terminator      805..809
                     /label=Transcriptional Termination 3
     misc_feature    813..817
                     /label=Transcriptional Termination 4
     terminator      829..833
                     /label=Transcriptional Termination 5
```

terminator 842..851
/label=Transcriptional Termination 6
BASE COUNT 264 a 226 c 250 g 281 t
ORIGIN
1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc
61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc
121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc
181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg
241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc
301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa
361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct
421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatatgc *ttaccgtaac*
481 *ttgaaagtat* ttcgatttct tggctttata tatcttgtgg aaaggacgaa acaccgataa
541 cttcgtatag catacattat acgaagttat tacacccgag ggggatgata aaccgggcgc
601 ggtcggtaaa gttgttccat tttttgaagc gaaggttgtg gatctggata ccgggaaaac
661 gctgggcgtt aatcaaagag gcgaactgtg tgtgagaggt cctatgatta tgtccggtta
721 tgtaaacaat ccggaagcgc cgcggccgct aggcaaggat gacacgcaaa ttcgtgaagc
781 gttccatatt tttacatcag gttgtttttc tgtttttaca tcaggttgtt tttctgtttg
841 gttttttttt tacaccacgt ttatacgccg gtgcacggtt taccactgaa aacacctttc
901 atctataact tcgtatagca tacattatac gaagttatcg acctcccgga ttccaattca
961 gcgggagcca cctgatttgg atcgggtggc tctcgctgag ttggaatcca tttttgaatt
1021 c (SEQ ID NO:27)

The PSE, the LoxP sites, and the transcription start site are marked as in the sequence shown for construct F. After recombination, only one LoxP site remains

and the sequence is the same as if the G nucleotide at position 536 is joined immediately 5' to the A nucleotide at position 905.

The promoter sequence of construct H is shown below.

```
LOCUS       U6 Lox Stop Lox 2A FF1   1012 bp    DNA
FEATURES             Location/Qualifiers
     misc_feature    58..513
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    536..536
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    470..489
                     /label=Proximal Sequence Element
     tRNA            505..513
                     /label=TATA Box
     misc_recomb     528..561
                     /label=LoxP
     misc_recomb     896..929
                     /label=LoxP
     terminator      611..616
                     /label=Transcriptional Termination 1
     misc_feature    562..730
                     /label=pGL3 Luciferase fragment
     terminator      744..895
                     /label=U6 Transcriptional Termination
     precursor_RNA   939..1001
                     /label=FF1 shRNA
     terminator      1002..1006
                     /label=Transcriptional Termination 7
     misc_feature    1007..1012
                     /label=EcoRI
     misc_feature    1..6
                     /label=XhoI
     terminator      780..784
                     /label=Transcriptional Termination 2
     terminator      796..800
                     /label=Transcriptional Termination 3
     misc_feature    804..808
                     /label=Transcriptional Termination 4
     terminator      820..824
                     /label=Transcriptional Termination 5
```

terminator 833..842
/label=Transcriptional Termination 6
BASE COUNT 261 a 220 c 251 g 280 t
ORIGIN

1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc

61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc

121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc

181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg

241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc

301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa

361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct

421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatatgc *ttaccgtaac*

481 *ttgaaagtat* ttcgatttct tggctttata tatcttgtgg aaaggacata **acttcgtata**

541 **gcatacatta tacgaagtta** ttacacccga gggggatgat aaaccgggcg cggtcggtaa

601 agttgttcca ttttttgaag cgaaggttgt ggatctggat accgggaaaa cgctgggcgt

661 taatcaaaga ggcgaactgt gtgtgagagg tcctatgatt atgtccggtt atgtaaacaa

721 tccggaagcg ccgcggccgc taggcaagga tgacacgcaa attcgtgaag cgttccatat

781 ttttacatca ggttgttttt ctgtttttac atcaggttgt ttttctgttt ggttttttt

841 ttacaccacg tttatacgcc ggtgcacggt ttaccactga aaacaccttt catctataac

901 **ttcgtatagc atacattata cgaagttatc** gacgaggcgg attccaattc agcgggagcc

961 acctgatttg gatcgggtgg ctctcgctga gttggaatcc atttttgaat tc (SEQ ID NO:28)

The PSE, the LoxP sites, and the transcription start site are marked as in the sequence shown for construct F. After recombination, the one of the LoxP site remains and the sequence is the same as if the T nucleotide at position 561 is joined immediately 5' to the C nucleotide at position 930.

The promoter sequence of construct I is shown below.

```
LOCUS       U6 Lox Stop Lox 2B   1012 bp    DNA
FEATURES             Location/Qualifiers
     misc_feature    58..513
                     /label=U6 Promoter
     promoter        8..25
                     /label=SP6 Promoter
     misc_feature    536..536
                     /label=Start of Transription
     enhancer        293..324
                     /label=Distal Sequence Element
     misc_feature    470..489
                     /label=Proximal Sequence Element
     tRNA            505..513
                     /label=TATA Box
     misc_recomb     528..561
                     /label=LoxP
     misc_recomb     896..929
                     /label=LoxP
     terminator      611..616
                     /label=Transcriptional Termination 1
     misc_feature    562..730
                     /label=pGL3 Luciferase fragment
     terminator      744..895
                     /label=U6 Transcriptional Termination
     precursor_RNA   939..1001
                     /label=FF1 shRNA
     terminator      1002..1006
                     /label=Transcriptional Termination 7
     misc_feature    1007..1012
                     /label=EcoRI
     misc_feature    1..6
                     /label=XhoI
     terminator      780..784
                     /label=Transcriptional Termination 2
     terminator      796..800
                     /label=Transcriptional Termination 3
     misc_feature    804..808
                     /label=Transcriptional Termination 4
     terminator      820..824
                     /label=Transcriptional Termination 5
     terminator      833..842
                     /label=Transcriptional Termination 6
```

BASE COUNT 260 a 223 c 248 g 281 t
ORIGIN
1 ctcgaggatt taggtgacac tatagaatac aagcttggct gcaggtcgac ggatcccccc
61 gagtccaaca cccgtgggaa tcccatgggc accatggccc ctcgctccaa aaatgctttc
121 gcgtcgcgca gacactgctc ggtagtttcg gggatcagcg tttgagtaag agcccgcgtc
181 tgaaccctcc gcgccgcccc ggccccagtg gaaagacgcg caggcaaaac gcaccacgtg
241 acggagcgtg accgcgcgcc gagcgcgcgc caaggtcggg caggaagagg gcctatttcc
301 catgattcct tcatatttgc atatacgata caaggctgtt agagagataa ttagaattaa
361 tttgactgta aacacaaaga tattagtaca aaatacgtga cgtagaaagt aataatttct
421 tgggtagttt gcagttttaa aattatgttt taaaatggac tatcatatgc *ttaccgtaac*
481 *ttgaaagtat* ttcgatttct tggctttata tatcttgtgg aaaggacata **acttcgtata**
541 **gcatacatta tacgaagtta** ttacacccga gggggatgat aaaccgggcg cggtcggtaa
601 agttgttcca ttttttgaag cgaaggttgt ggatctggat accgggaaaa cgctgggcgt
661 taatcaaaga ggcgaactgt gtgtgagagg tcctatgatt atgtccggtt atgtaaacaa
721 tccggaagcg ccgcggccgc taggcaagga tgacacgcaa attcgtgaag cgttccatat
781 ttttacatca ggttgttttt ctgtttttac atcaggttgt ttttctgttt ggtttttttt
841 ttacaccacg tttatacgcc ggtgcacggt ttaccactga aaacaccttt catctataac
901 **ttcgtatagc atacattata cgaagttatc** gacctcccgg attccaattc agcgggagcc
961 acctgatttg gatcgggtgg ctctcgctga gttggaatcc atttttgaat tc (SEQ ID NO:29)

The PSE, the LoxP sites, and the transcription start site are marked as in the sequence shown for construct F. After recombination, one of the LoxP sites remains and the sequence is the same as if the T nucleotide at position 561 is joined immediately 5' to the C nucleotide at position 930.

Constructs F-I were made by introducing the FF1 sequence and the LoxP Stop cassette into the human U6 promoter DNA sequence using PCR. The

U6 promoter DNA fragments were then digested with EcoRI and XhoI and inserted into pENTR11 (Invitrogen) digested with these two enzymes. The constructs were cotransfected into NIH 3T3 cells with pGL3 and pSEAP2. Forty-eight hours after the start of transfection, 10 μl supernatant from the cells was removed and used for a SEAP luminescence assay (see above). The cells were subjected to a Luciferase luminescence assay (Steady-Glo, Promega). Data were normalized by dividing Luciferase assay values with SEAP assay values. The comparison between the empty vector (pENTR11) and U6 FF1 vectors shows the degree of inhibition of Luciferase expression from pGL3.

Fig. 3B shows that the recombined forms of construct F-I effectively silenced Luciferase expression. The non-recombined forms of these constructs, in contrast, showed no silencing of Luciferase.

In conclusion, Cre-mediated removal of the LoxP Stop cassette allows effective expression of shRNAs from the U6 promoter. These shRNAs are capable of efficient gene-specific silencing even though the LoxP sequence is also present in the transcript. These vectors will allow one to temporally and spatially restrict the expression of shRNAs through the use of spatially and/or temporally-restricted promoters. Additionally, the U6 LoxP Stop vector design can be combined with IPTG or Tetracycline inducible U6 promoters (below) to allow inducible control after Cre-mediated removal of the LoxP Stop. This allows the advantages of inducible expression of shRNAs while restricting expression only to specific tissues, thus preventing possible systemic effects of expression of shRNAs in non-target tissues.

C. Tetracycline-Inducible Promoters

The Tet operon contains two different operator sequences, designated TO1 (5'-ACTCTATCATTGATAGAGT-3'; SEQ ID NO:30) and TO2 (5'-TCCCTATCAGTGATAGAGA-3'; SEQ ID NO:31), respectively. Both TO1 and TO2 bind the tetracycline repressor (TetR) protein. TO1 and TO2 sequences were inserted between the TATA and the transcription initiation site in the human U6 promoter at various positions while maintaining the spacing between the TATA and the initiation site and preserving the C preceding the G at the initiation site. These constructs are designated J, K, L, M, N, O, and P (also referred to as

U6TO1A, U6TO1B, U6TO1C, U6TO2A, U6TO2B, U6TO2C, and U6TO2D, respectively). Their sequence alignment (from the PSE to the transcriptional start) is shown below (SEQ ID NOs:32-38).

```
J cttaccgtaacttgaaagtatttcgatttcttggctttatatatACTCTATCATTGATAGAGTaccg
K cttaccgtaacttgaaagtatttcgatttcttggctttatatataACTCTATCATTGATAGAGTccg
L cttaccgtaacttgaaagtatttcgatttcttggctttatatataaACTCTATCATTGATAGAGTcg
M cttaccgtaacttgaaagtatttcgatttcttggctttatatatCCCTATCAGTGATAGAGAcaccg
N cttaccgtaacttgaaagtatttcgatttcttggctttatatatTCCCTATCAGTGATAGAGAaccg
O cttaccgtaacttgaaagtatttcgatttcttggctttatatatcTCCCTATCAGTGATAGAGAccg
P cttaccgtaacttgaaagtatttcgatttcttggctttatatatcttCCCTATCAGTGATAGAGAcg
```

In the above alignment, the PSE is italicized; the TATA box is shown in a box; the TO1 and TO2 sequences are capitilized; and the first G of the transcription initiation site is double-underlined.

Constructs J-P were constructed as described below:

To insert the TetOp sequence into the promoter region, the following primer sets were used in two PCRs: (1) a first external primer SEQ ID NO:17 (*supra*), which targeted a vector region 5' to the U6 transcription unit, and a first internal primer having the following sequence for each construct of:

J: 5'-agcacggtactctatcaatgatagagtatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:39),

K: 5'-agcacggactctatcaatgatagagttatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:40),

L: 5'-agcacgactctatcaatgatagagtttatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:41,

M: 5'-agcacggtgtctctatcactgatagggatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:42),

N: 5'-agcacggttctctatcactgatagggaatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:43),

O: 5'-agcacggtctctatcactgatagggagatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:44), and

P: 5'-agcacgtctctatcactgatagggaagatatataaagccaagaaatcgaaatact-3' (SEQ ID NO:45);

and (2) a second internal primer having following sequence for each construct of:

J: 5'-tggctttatatatactctatcattgatagagtaccgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:46),

K: 5'-tggctttatatataactctatcattgatagagtccgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:47),

L: 5'-tggctttatatataaactctatcattgatagagtcgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:48),

M: 5'-tggctttatatatccctatcagtgatagagacaccgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:49),

N: 5'-tggctttatatattccctatcagtgatagagaaccgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:50),

O: 5'-tggctttatatatctccctatcagtgatagagaccgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:51), and

P: 5'-tggctttatatatcttccctatcagtgatagagacgtgctcgcttcggcagcacatatac-3' (SEQ ID NO:52);

and a second external primer that targeted a vector region 3' to the U6 transcription unit. The PCR products from these two reactions were mixed and subjected to PCR again using the two external primers. This PCR reaction generated a complete, modified U6 transcription unit having the TetOp sequence. To construct an RNAi vector, the U6-coding sequence was then replaced in part or in its entirety by a sequence encoding an interfering dsRNA, using routine recombinant techniques (e.g., PCR).

U6 promoter DNA fragments containing the FF1 sequence were digested with EcoRI and XhoI and inserted into pENTR11 (Invitrogen) digested with the same enzymes. U6 promoter constructs were cotransfected with pGL3 Control (Luciferase expression, Promega), pcDNA6/TR (for directing TetR expression, Invitrogen), and pSEAP2 Control (SEAP expression used as a cotransfection control, BD Biosciences) into NIH 3T3 cells. Cells were treated with tetracycline and were compared to untreated cells. Forty-eight hours after the start of transfection, 10 μl supernatant from the cells was removed and used for a SEAP luminescence assay. The cells were subjected to a Luciferase luminescence assay (Steady-Glo, Promega). Data were normalized by dividing Luciferase assay

values with SEAP assay values. These values were then normalized to an empty vector control.

As shown in Fig. 4A, the TetO constructs maintained wildtype transcriptional activity in the absence of TetR. When TetR and tetracycline were present, the constructs significantly inhibited Luciferase expression, but this inhibition was significantly repressed by TetR in the absence of tetracycline (Fig. 4B). Construct J consistently appeared to be the most strongly repressed U6 TO1 promoter, and construct N consistently appeared to be the most strongly repressed U6 TO2 promoter. Notably, the positions of the TetOp in constructs J and N were the same. In conclusion, when inserting the TetOp sequence into the U6 promoter, it is important to maintain the spacing between the TATA box and the transcription initiation site, and to preserve the C nucleotide immediately 5' to the G initiation site.

Because the TetR gene is a prokaryotic gene, we designed a version of the gene that is codon-optimized for expression in mammalian cells. This version was designated gpTetR. This sequence was also designed to prevent possible GpC methylation, which would prevent silencing of the gene in mammals. An SV40 nuclear localization signal was added to the C-terminal end to allow localization of the protein product to the nucleus, allowing gpTetR to bind the TetOp more efficiently.

The gpTetR sequence is as follows.

```
LOCUS       gpTetR         648 bp    DNA
FEATURES             Location/Qualifiers
misc_signal     622..645
                /label=SV40 Nuclear Localization Signal
BASE COUNT      190 a       149 c       164 g
145 t
  1 atgagcaggc ttgacaaatc aaaagtgatt aactcagctc ttgaattgct
caatgaagtg
 61 gggatcgagg gtctaactac acgaaagctg gcacagaagc taggggtgga
acagccaacc
121 ctgtattggc atgtgaaaaa caaaagagcc ctgcttgatg cactggctat
tgagatgcta
```

181 gacagacacc atacccactt ctgtcctctg gaaggggaga gctggcagga cttcctgaga

241 aacaatgcga agtctttccg ttgtgcactc ctgagccatc gcgatggagc caaagttcat

301 ttagggactc ggcccacaga aaagcaatac gagacactag agaatcagct cgcctttctg

361 tgccagcaag gctttagtct ggaaaatgca ctctatgctc tcagcgccgt gggacacttt

421 accttgggat gtgtccttga agatcaagag catcaggttg caaaggaaga gagagaaact

481 ccaactacag acagtatgcc cccattgctg aggcaggcta tagaattatt cgaccaccag

541 ggcgcagaac ctgcctttct ctttggtctg gagctgatta tttgtggctt agagaaacaa

601 ctcaaatgtg aatcaggctc tccacctaag aagaaacgga aggtttaa (SEQ ID NO:53)

The inducible tetracycline system was further validated using stable cell lines and gpTetR. Human colorectal carcinoma cells, HCT116, were infected with virus containing a luciferase expression construct containing the Hygromycin resistance gene. After selection with Hygromycin, the cells were subsequently transfected with gpTetR, under the control of the PGK promoter. The gpTetR expression construct contains a Zeocin resistance gene. The stably transfected cells were selected with Zeocin and subcloned. Single clones were tested for adequate gpTetR expression, using real time RT-PCR, and subsequently infected with lentiviral vectors containing either the luciferase-directed FF1 shRNA under the control of the inducible promoter U6TO2B (promoter construct N), or empty vector. As a control to test the effect of the hairpin construct in the absence of the gpTetR repressor, HCT116 cells were infected with virus containing the luciferase expression construct and then infected with the empty virus or U6TO2B-FF1 lentivirus. 1 ug/ml Doxycyline was added to the cells and luciferase activity was measured after 72 hours. The luciferase values were determined by treating cells with Steady Glo reagent (Promega) and measuring luminescence with a plate reader (Bio-Tek). Luciferase values were averaged from 4 replicates for each sample. The standard deviation was then calculated from those replicates. Values

and standard deviations were rescaled with empty vector equal to 1. Figure 5A shows that in cells lacking gpTetR, FF1-driven RNAi of luciferase occurs with or without Doxycycline. On the other hand, cells containing gpTetR show repression of RNAi in the absence of Doxycyclin, and FF1-driven RNAi of luciferase in the presence of Doxycycline.

The system was then tested *in vivo*. Cell were produced as described above. Nude mice were injected sub-cutaneously with 10^6 cells per injection site, at three separate sites, on both flanks. Doxycyline (2 μg/ml)was admininstered orally in drinking water (changed twice a week). Tumor size and luciferase activity were measured over a period of 14 days, using a caliper and the NightOWL imaging system (Berthold Technolgies GmbH & Co KG), respectively. The luciferase values for each tumor at each time point were determined by injecting (interperitoneal) mice with luciferin (Molecular Probes) and measuring luminescence with the NightOwl imaging system. Each tumor luminescence value was divided by its size (area), and then these values were averaged per sample type. Average values were rescaled so that the average value for tumors not containing U6TO2B FF1 (thus containing empty virus control, not shown in Figure 5B) is equal to one for each time point. HCT116 Luciferase PGKgpTetR (empty virus control) is revalued to one when comparing to HCT116 Luciferase PGKgpTetR U6TO2B FF1 (minus doxycycline samples are compared to each other and plus doxycycline samples are compared to each other). HCT116 Luciferase (empty virus control) is revalued to one when comparing to HCT116 U6TO2B FF1 (both plus and minus doxcycline samples are grouped together for each tumor type). The results show that the down-regulation, repression, and de-repression phenotypes seen *in vitro* are also observed *in vivo* (Fig. 5B). In summary, gpTetR effectively represses the U6TO2B promoter (construct N) in the absence of Doxycycline, and the presence of Doxycycline allows effective expression from U6TO2B, allowing RNAi to function.

D. Combined Promoter Systems and an Alternative Use for the Cre-Lox System

This example describes a promoter system combining a small molecule inducible system (e.g., the tetracycline-inducible system or the IPTG-

inducible system) with the Cre-Lox system (Fig. 6). By combining the two systems, one can control RNAi both spatially and temporally. First, a small molecule inducible U6 promoter, containing a LacO or TetOp sequence, is fused to a LoxP-Stop-LoxP shRNA cassette. The LoxP-Stop-LoxP cassette prevents transcription of the shRNA, preventing RNAi of a target gene. Upon Cre Recombinase-mediated recombination, the Stop cassette is removed, leaving only one LoxP site (Fig. 6). Tissue-specific expression of Cre Recombinase allows spatial control of shRNA expression. Temporal control can be achieved by a promoter that is temporally restricted during development or through an inducible form of Cre. An example of inducible Cre is a fusion between Cre and the Estrogen Receptor (ER). Tamoxifen allows the Cre-ER fusion to enter the nucleus and recombine the LoxP sites, removing the LoxP Stop cassette. Temporal control of Cre can also be accomplished by the use of an inducible promoter for Cre.

After Cre-mediated removal of the LoxP-Stop-LoxP cassette, expression of the shRNA from the U6 promoter can further be controlled by the use of U6 promoters containing Lac Operator or Tet Operator sequences (Fig. 6). When LacI is also expressed in the cell, expression from the U6 Lac Operator promoter will only occur in the presence of IPTG. When TetR is also expressed in cells, expression from the U6 Tet Operator promoter will only occur in the presence of tetracycline or doxycycline. The addition of the Lac Operator or Tet Operator controlled system to the Cre-Lox controlled system allows one to control the amount of expression from the U6 promoter in specific tissues and at desired times. The Cre-Lox system alone allows one to turn on shRNA expression at desired times in desired tissues. However, once the LoxP Stop LoxP cassette is removed, transcription of the shRNA is continuous. The addition of the Lac Operator or Tet Operator system allows one to turn on and off expression of the shRNA at desired times in desired tissues. Additionally, the level of shRNA expression can be controlled by using different doses of the small molecule inducer.

An alternative use of the Cre-Lox system is to use Cre-mediated recombination to switch from expression of one shRNA to another (Fig. 7). Instead of using a Stop sequence between two LoxP sites, a shRNA is placed

immediately after the first LoxP sequence and is followed by a stretch of four or more Ts. The second LoxP sequence (in the same orientation as the first), is followed by a second shRNA. Before Cre-mediated recombination, the first shRNA will be expressed and transcription terminates before the second one. After Cre-mediated recombination, only the second shRNA is present and expressed.

What is claimed is:

1. A nucleic acid construct comprising a coding sequence for a small interfering RNA molecule, said coding sequence linked operably to a mammalian or viral promoter,

wherein a part of the nucleotide region between the PSE and the transcription initiation site of the promoter has been replaced with an operator sequence, wherein said operator sequence is controlled by a small molecule, and

wherein said operator sequence has the same length as said part, or is no more than two nucleotides longer than said part.

2. The nucleic acid construct of claim 1, wherein said promoter is an RNA pol I, pol II, or pol III promoter.

3. The nucleic acid construct of claim 2, wherein said promoter is a U6, H1 or CMV promoter.

4. The nucleic acid construct of claim 1, wherein said operator sequence is between the TATA box and the transcription initiation site of the promoter.

5. The nucleic acid construct of claim 1, wherein said operator sequence is a Lac operator sequence or a tetracycline operator sequence.

6. The nucleic acid construct of claim 1, wherein said Lac operator sequence is SEQ ID NO:2 or 4, or said sequence with up to 4 terminal deletions from either end of the sequence.

7. The nucleic acid construct of claim 5, selected from SEQ ID NO: 6 or 7.

8. The nucleic acid construct of claim 1, wherein a part of the nucleotide region between the TATA box and the transcription initiation site of the promoter has been replaced with a first Lac operator sequence, and wherein a part of the nucleotide region between the distal element sequence and the proximal element sequence has been replaced with an additional Lac operator sequence.

9. The nucleic acid construct according to claim 8, wherein the first LacO sequence and the additional LacO sequence are selected from the group consisting of SEQ ID NOS: 2-5.

10. The nucleic acid construct according to claim 9, selected from the group consisting of: SEQ ID NOS: 8-10.

11. The nucleic acid construct of claim 5, wherein said tetracycline operator sequence is SEQ ID NO:30 or 31.

12. The nucleic acid construct according to claim 11, selected from the group consisting of: SEQ ID NOS: 32-38.

13. The nucleic acid construct of claim 5, wherein said tetracycline operator sequence is placed immediately downstream from the TATA box.

14. The nucleic acid construct of claim 1, wherein said small interfering RNA molecule inhibits expression of a disease-related gene.

15. The nucleic acid construct of claim 14, wherein the disease is cancer.

16. A nucleic acid construct comprising a coding sequence for a small interfering RNA molecule linked operably to a mammalian or viral promoter,

wherein the transcription initiation site of the promoter and said coding sequence are separated by a LoxP-Stop-LoxP cassette consisting of a transcription termination sequence flanked by two LoxP sites,

wherein upon recombination between said two LoxP sites in the presence of a cre recombinase, said transcription termination sequence is removed, allowing transcription to proceed through said coding sequence.

17. The nucleic acid construct of claim 16, wherein said promoter is an RNA pol I, pol II, or pol III promoter.

18. The nucleic acid construct of claim 17, wherein said promoter is a U6, H1 or CMV promoter.

19. The nucleic acid construct according to claim 18, selected from the group consisting of SEQ ID NOS: 26-29.

20. The nucleic acid construct of claim 16, wherein said small interfering RNA molecule inhibits expression of a disease-related gene.

21. The nucleic acid construct of claim 20, wherein the disease is cancer.

22. A nucleic acid construct comprising a coding sequence for a first small interfering RNA molecule linked operably to a mammalian RNA polymerase III promoter,

wherein the first small interfering RNA molecule is placed between a LoxP site and a stretch of at least four thymines,

wherein the construct further comprises a second LoxP site followed by a coding sequence for a second small interfering RNA molecule,

wherein before recombination between said two LoxP sites in the presence of cre recombinase, the first small interfering RNA molecule is expressed; and

wherein upon recombination between said two LoxP sites in the presence of a cre recombinase, the first small interfering RNA molecule is no longer expressed and the second small interfering RNA molecule is expressed.

23. The nucleic acid construct of claim 16 or claim 22, further comprising an operator sequence, wherein said operator sequence is controlled by a small molecule, and wherein said operator sequence replaces a portion of nucleic acid sequence of a selected from: a portion between the TATA box and the transcription initiation site, a portion between the PSE and the TATA box, a portion between the distal element sequence and the proximal element sequence and both a portion between the TATA box and the transcription initiation site and a portion between the distal element sequence and the proximal element sequence.

24. A mammalian cell comprising the nucleic acid construct of any one of claims 1 to 15 and claim 23.

25. The mammalian cell according to claim 24, further comprising an additional construct for the expression of a repressor of the operator sequence.

26. The mammalian cell according to claim 25 , wherein the additional construct encodes LacI or TetR.

27. A mammalian cell comprising the nucleic acid construct according to any one of claim 16 to 22.

28. The mammalian cell according to claim 27, further comprising a construct that encodes Cre recombinase.

29. The cell according to any one of claims 25 or 26, wherein expression of the repressor is under the control of a promoter selected from the

group consisting of: an inducible promoter, a stage-specific promoter and a tissue-specific promoter.

30. The cell according to claim 28, wherein Cre recombinase expression is under the control of a promoter selected from the group consisting of: an inducible promoter, a stage-specific promoter and a tissue-specific promoter.

31. A non-human mammal comprising the cell according to any one of claims 24 to 30.

32. The non-human mammal according to claim 31, which is a chimeric mammal some of whose somatic or germ cells are a cell according to any one of claims 24 to 30.

33. The non-human mammal according to claim 31, which is a transgenic mammal all of whose somatic or germ cells are cells according to any one of claims 24 to 30.

34. A method for making a chimeric non-human mammal according to claim 32 comprising the step of introducing a construct according to any one of claims 1-23 into an embryonic stem (ES) cell and generating a chimeric mammal from the ES cell.

35. A method for making a transgenic non-human mammal according to claim 33 comprising the step of mating a chimeric non-human mammal according to claim 34 with another animal from the same species.

36. A tetracycline repressor encoded by SEQ ID NO:53 or a degenerate variant thereof.

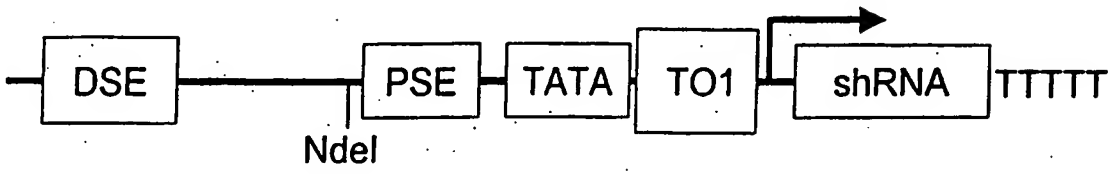
37. A method for inhibiting the expression of a gene of interest in a cell comprising the step of introducing a construct according to any one of

claims 1 to 23 into the cell, wherein the small interfering RNA molecule is specific for the gene of interest.

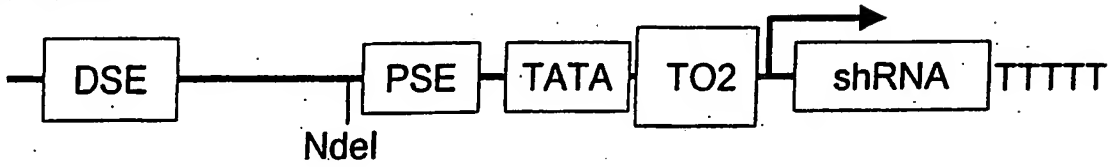
38. A method for treating a gene-mediated disease comprising introducing into a cell a construct according to any one of claims 1 to 23 where the small interfering RNA molecule is specific for the gene mediating the disease.

Figure 1

J, K, L



M, N, O, P



DSE = Distal Sequence Element
TATA = TATA Box
shRNA = short hairpin RNA

PSE = Proximal Sequence Element
TO = Tet Operator 1 or 2

Figure 2A. Inducible U6 Lac Operator Vectors
A
DSE
PSE
NdeI
TATA
Natural LO
shRNA
TTTTT
B
DSE
PSE
NdeI
TATA
Synthetic LO
shRNA
TTTTT
C
DSE
Synthetic LO
PSE
TATA
Natural LO
shRNA
TTTTT
D
DSE
Natural LO
PSE
TATA
Natural LO
shRNA
TTTTT
E
DSE
Synthetic LO
PSE
TATA
Synthetic LO
shRNA
TTTTT
DSE = Distal Sequence Element
TATA = TATA Box
Synthetic LO = Synthetic Lac Operator (U6 LOP)
shRNA = short hairpin RNA
PSE = Proximal Sequence Element
Natural LO = Natural Lac Operator

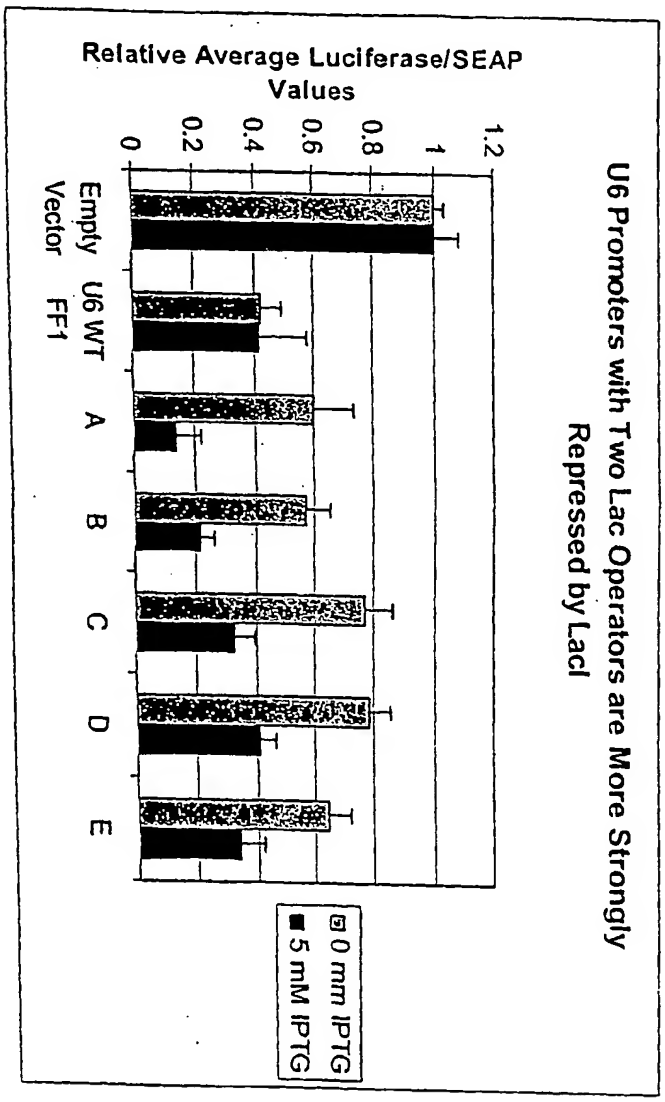

U6 Promoters with Two Lac Operators are More Strongly Repressed by LacI
Relative Average Luciferase/SEAP Values
0
0.2
0.4
0.6
0.8
1
1.2
Empty Vector
U6 WT
FF1
A
B
C
D
E
0 mM IPTG
5 mM IPTG

Figure 2B

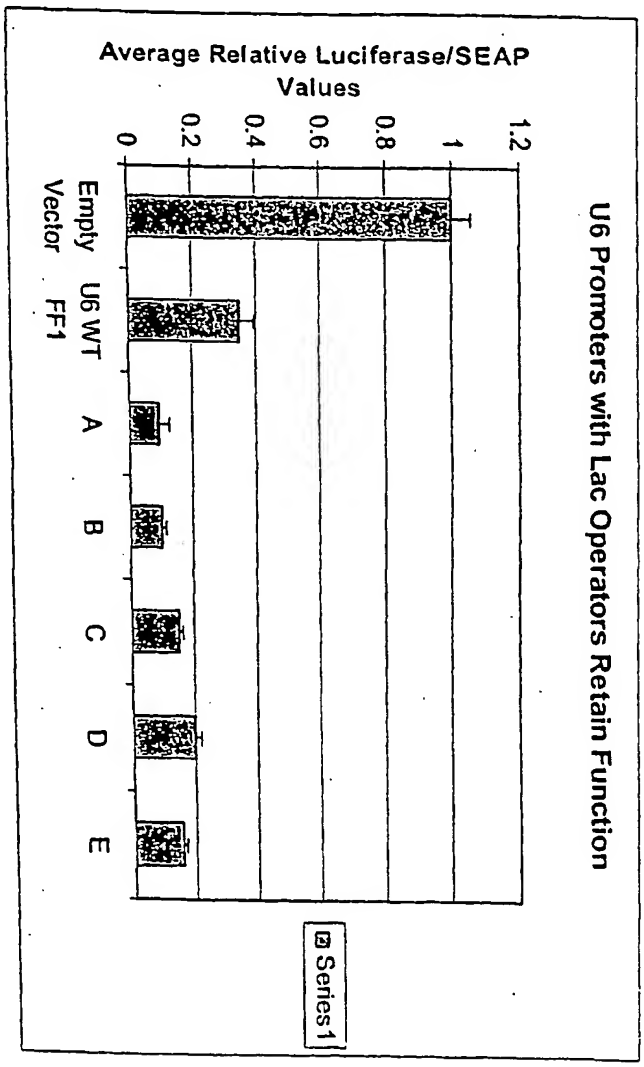

U6 Promoters with Lac Operators Retain Function
Average Relative Luciferase/SEAP Values
0
0.2
0.4
0.6
0.8
1
1.2
Empty Vector
U6 WT
FF1
A
B
C
D
E
Series1

Figure 2C

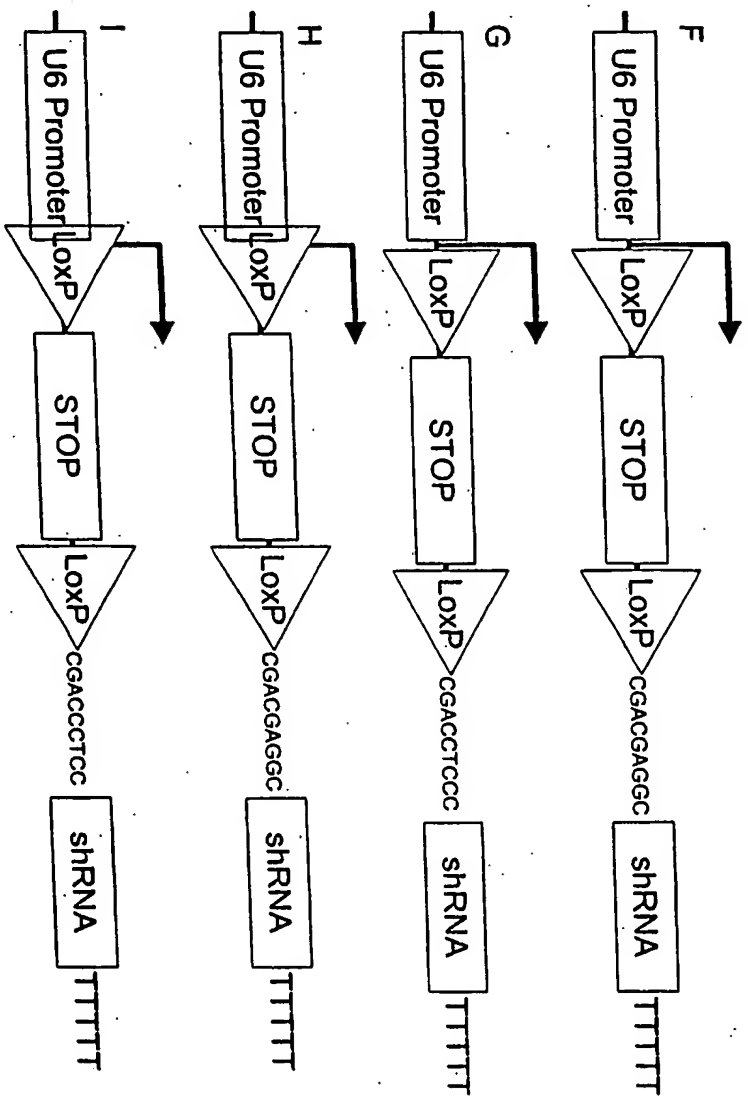

F
U6 Promoter
LoxP
STOP
LoxP
CGACGAGGC
shRNA
TTTTT
G
U6 Promoter
LoxP
STOP
LoxP
CGACCTCCC
shRNA
TTTTT
H
U6 Promoter
LoxP
STOP
LoxP
CGACGAGGC
shRNA
TTTTT
I
U6 Promoter
LoxP
STOP
LoxP
CGACCCTCC
shRNA
TTTTT

Figure 3A

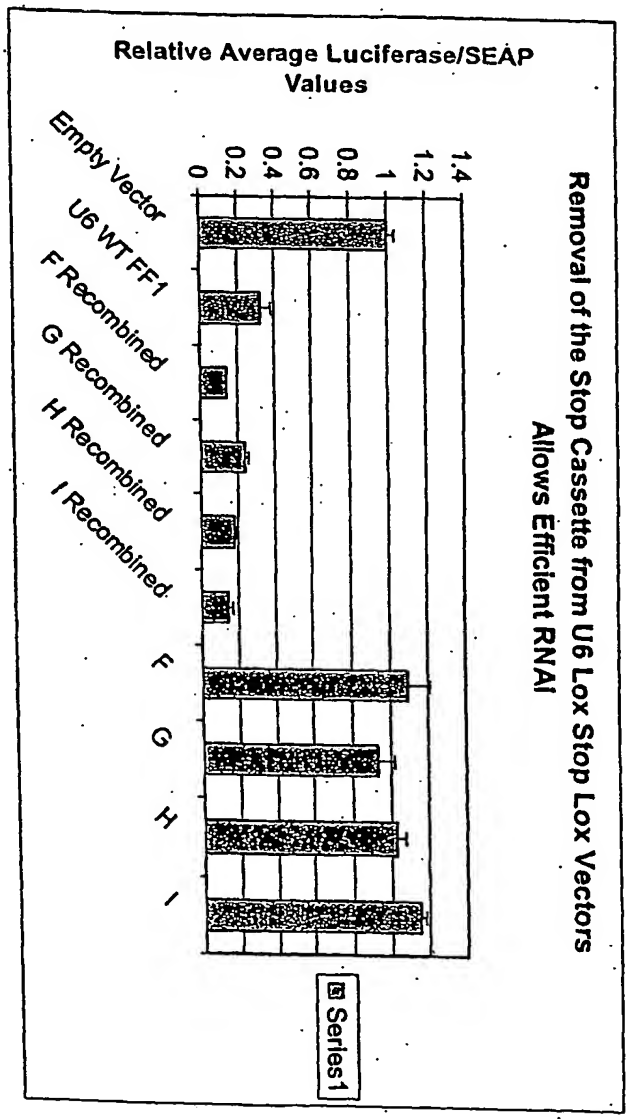

Removal of the Stop Cassette from U6 Lox Stop Lox Vectors Allows Efficient RNAi
Relative Average Luciferase/SEAP Values
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Empty Vector
U6 WT FF1
F Recombined
G Recombined
H Recombined
I Recombined
F
G
H
I
Series1

Figure 3B

FIG. 3C

FIG. 3D

FIG. 3E

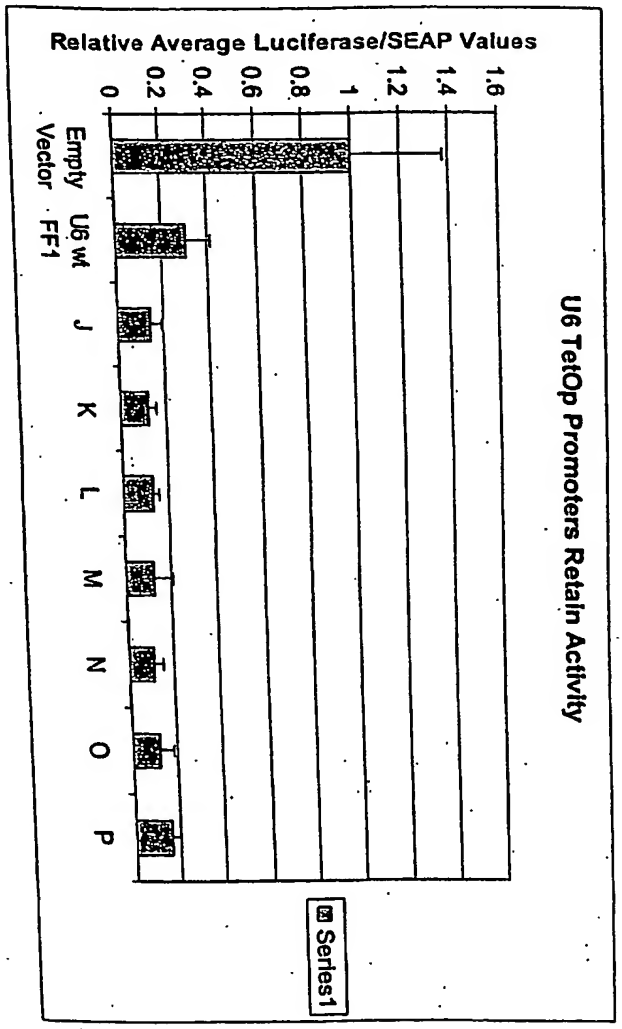

Relative Average Luciferase/SEAP Values
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
Empty Vector
U6 wt FF1
J
K
L
M
N
O
P
U6 TetOp Promoters Retain Activity
Series1

Figure 4A

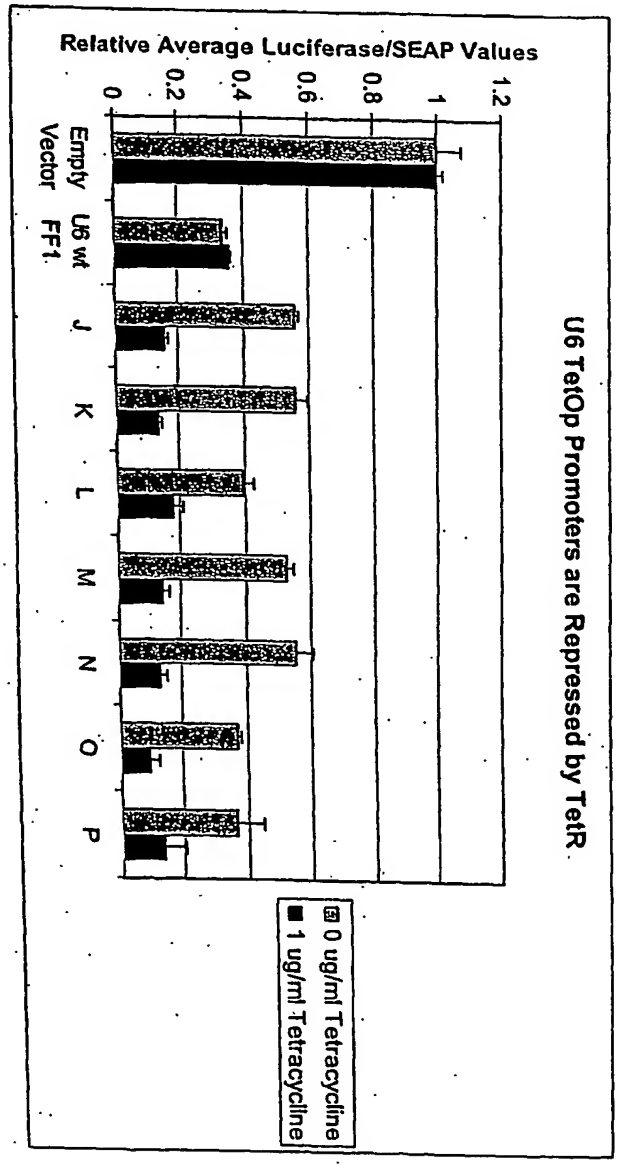

Relative Average Luciferase/SEAP Values
0
0.2
0.4
0.6
0.8
1
1.2
Empty Vector
U6 wt FF1
J
K
L
M
N
O
P
U6 TetOp Promoters are Repressed by TetR
0 ug/ml Tetracycline
1 ug/ml Tetracycline

Figure 4B

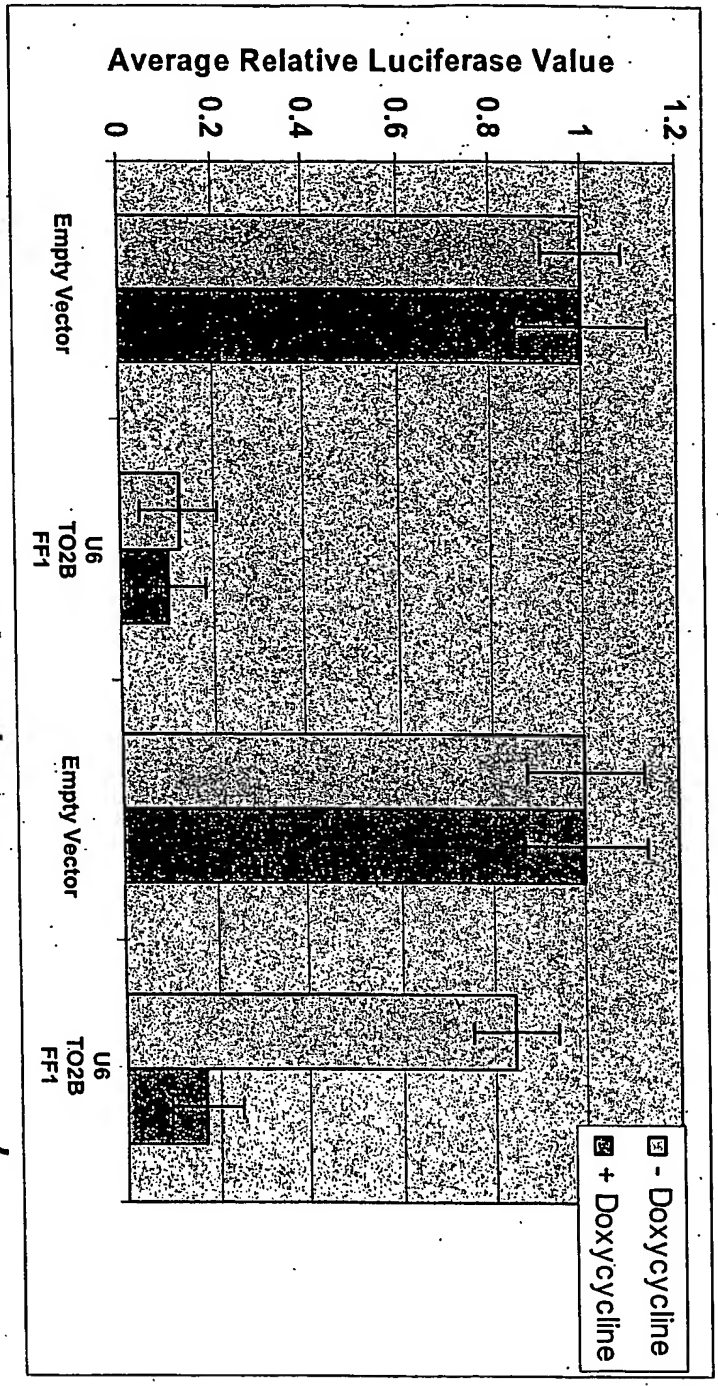

Average Relative Luciferase Value
0
0.2
0.4
0.6
0.8
1
1.2
- Doxycycline
+ Doxycycline
Empty Vector
U6 TO2B FF1
Empty Vector
U6 TO2B FF1
HCT116-Luc
HCT116-Luc-PGKgpTetR

Figure 5A

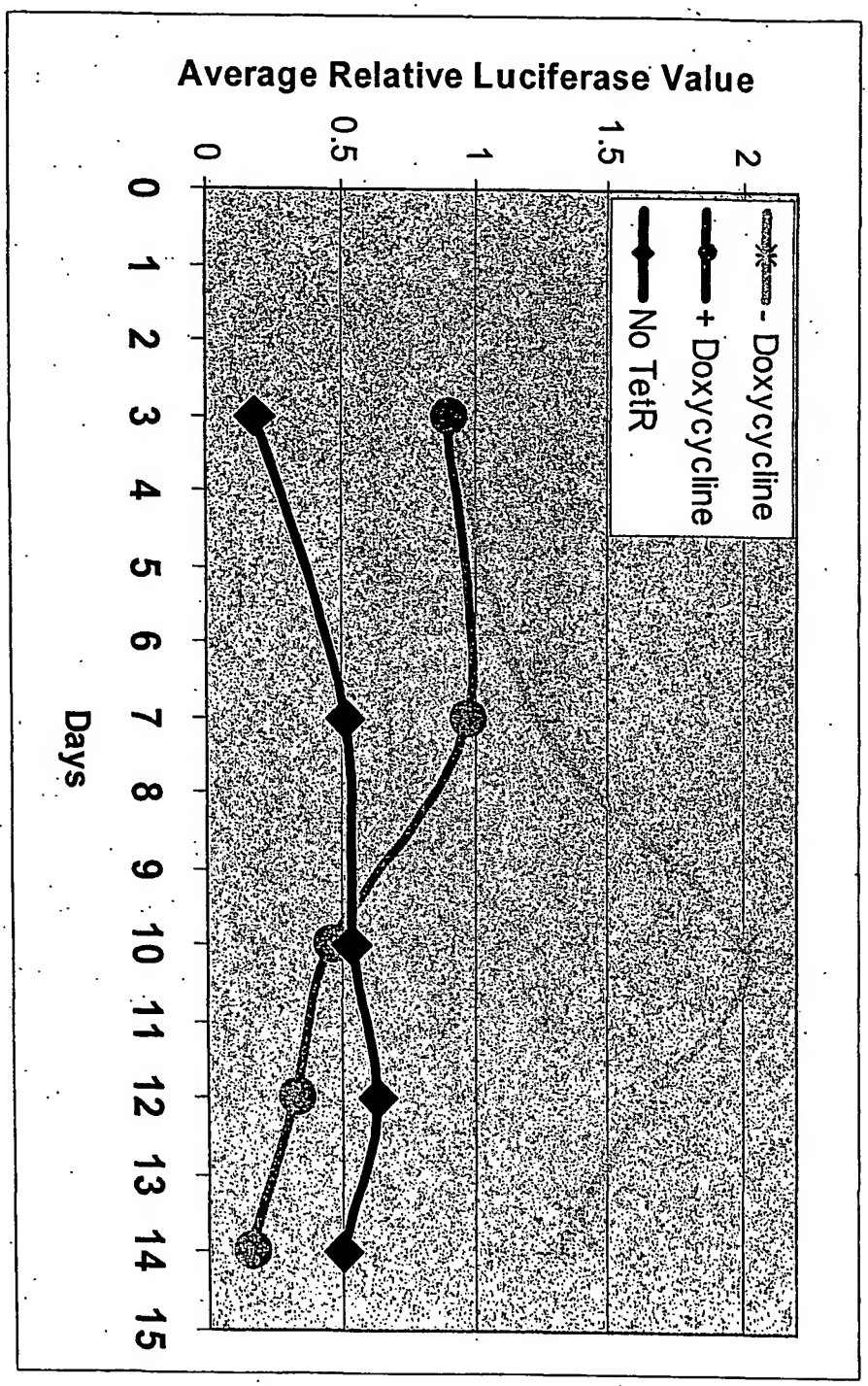

Average Relative Luciferase Value
0
0.5
1
1.5
2
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Days
- Doxycycline
+ Doxycycline
No TetR

Figure 5B

Figure 6

DSE
NdeI
PSE
TATA
TO or LO
LoxP
STOP
LoxP
shRNA
TTTTT
Cre Recombinase
DSE
NdeI
PSE
TATA
TO or LO
LoxP
shRNA
TTTTT

Figure 7

U6 Promoter
LoxP
shRNA1
TTTTT
LoxP
shRNA2
TTTTT
Cre Recombinase
U6 Promoter
LoxP
shRNA2
TTTTT

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 July 2004 (08.07.2004)

PCT

(10) International Publication Number
WO 2004/056964 A3

(51) International Patent Classification[7]: C12N 5/10, 15/09, 15/31, 15/85, C07K 14/24

(21) International Application Number: PCT/US2003/040548

(22) International Filing Date: 18 December 2003 (18.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/434,856 18 December 2002 (18.12.2002) US
60/499,313 28 August 2003 (28.08.2003) US

(71) Applicant *(for all designated States except US)*: **GENPATH PHARMACEUTICALS, INCORPORATED** [US/US]; 300 Technology Square, 7th Floor, Cambridge, MA 02139 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **WINSTON, William** [US/US]; 108 Liberty Avenue, Apt. #1, Somerville, MA 02144 (US). **O'HAGAN, Ronan, C.** [CA/US]; 15 Garrison Road, Apt. 2, Brookline, MA 02445 (US). **HEYER, Joerg** [DE/US]; 6 Fainwood Circle, Cambridge, MA 02139 (US). **RIDEOUT, William** [US/US]; 20 Corporal Burns Road, Cambridge, MA 02142 (US). **ETEMAD-MOGHADAM, Bijan** [US/US]; 289 Vermont Street, West Roxbury, MA 02132 (US).

(74) Agents: **GUNNISON, Jane, T.** et al.; c/o Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

(88) Date of publication of the international search report: 14 October 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: VECTORS FOR INDUCIBLE RNA INTERFERENCE

(57) Abstract: Recombinant vectors for inducibly expressing double-stranded RNA molecules that interfere with the expression of a target gene.

WO 2004/056964 A3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/40548

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12N 5/10, 15/09, 15/31, 15/85; C07K 14/24
US CL : 435/320.1, 325; 530/350
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/320.1, 325; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	HASUWA et al. Small interfering RNA and gene silencing in transgenic mice and rats. FEBS Letters. 13 November 2002, Vol. 532, pages 227-230, especially page 230, column 2.	1-5, 11, 13, 14
— A		1-23, 27, 28, 30
Y	HANNON, G. J. RNA interference. Nature. 11 July 2002, Vol. 418, pages 244-251, especially page 250, columns 1-2.	1-5, 11, 13-15
— A		1-23, 27, 28, 30
Y	ELBASHIR et al. Analysis of gene function in somatic mammalian cells using small interfering RNAs. Methods. February 2002, Vol. 26, No. 2, pages 199-213, especially page 201, col. 2.	1-5, 11, 13, 14
Y	OHKAWA et al. Control of the functional activity of an antisense RNA by a tetracycline-responsive derivative of the human U6 snRNA promoter. Human Gene Therapy. 01 March 2000, Vol. 11, No. 4, pages 577-585, especially page 578, Figure 1; page 579, Figure 2; page 581, Figure 3; and page 584.	1-5, 11, 13-15
Y	US 5,917,122 A (BYRNE, G.) 29 June 1999 (29.06.1999), columns 15-18, 39 and 40.	36

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
21 June 2004 (21.06.2004)

Date of mailing of the international search report
13 AUG 2004

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Authorized officer
Scott D. Priebe

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

PCT/US03/40548

INTERNATIONAL SEARCH REPORT

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	OGUETA et al. Design and in vitro characterization of a single regulatory module for efficient control of gene expression in both plasmid DNA and a self-inactivating lentiviral vector. Molecular Medicine. August 2001, Vol. 7, No. 8, pages 569-579, see especially page 569, column 1; pages 570 and 571; and page 575, col. 2 to page 576, col. 1.	36
E, X	US 2004/0005593 A1 (LORENS, J.) 08 January 2004 (08.01.2004), Figure 3, paragraphs 0018, 0060-0062, 0070-0074.	1-3, 5, 11
E, X	US 2004/0115815 A1 (LI et al.) 17 June 2004 (17.06.2004), Figure 6, paragraphs 0009, 0026, 0075-0077, 0136-0139, 0151, 0235-0236, 0266-0271.	1-3, 5, 11, 14, 15
E, X	US 2004/0002077 A1 (TAIRA et al.) 01 January 2004 (01.01.2004), Figure 12, paragraphs 0116, 0209-0211.	1-3, 5, 11
P, X	SASAKI et al. A system for conditional RNA interference in the mouse using the lac operator - repressor system. Society for Neuroscience Abstract Viewer and Itinerary Planner. 08 November 2003, Vol. 2003, Abstract No. 325.4.	1-3, 5, 6
P, X	CZAUDERNA et al. Inducible shRNA expression for application in a prostate cancer model. Nucleic Acids Research. 01 November 2003, Vol. 31, No. 21, e127, pages 1-7, see entire reference especially page 4, Figure 2.	1-5, 11, 13-15
A	SUI et al. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. Proceedings of the National Academy of Sciences USA. 16 April 2002, Vol. 99, No. 8, pages 5515-5520.	1-23, 27, 28, 30

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/40548

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claim Nos.: 24-26, 29, 31-35, 37, 38
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/US03/40548

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-15, drawn to a nucleic acid construct for expression of a siRNA comprising a promoter including an operator.

Group II, claim(s) 16-21, 23, 27, 28, and 30, drawn to a nucleic acid construct for expression of a siRNA comprising a LoxP-Stop-LoxP cassette between the transcription initiation site and the coding sequence for the siRNA.

Group III, claim(s) 22, 23, 27, 28, and 30, drawn to a nucleic acid construct for expression of two siRNAs comprising, in order, a pol III promoter, a first loxP sequence, coding sequence for a first siRNA, at least four thymidines, a second loxP sequence, and coding sequence for a second siRNA.

Group IV, claim(s) 36, drawn to a tetracycline repressor protein.

The inventions listed as Groups I-IV do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Groups I-III are each related to different constructs for expression of siRNA. The technical feature of a construct for expressing siRNA was known in the prior art, see Hasuwa et al. for example. Consequently, this technical feature shared by Groups I-III is not a special technical feature. The constructs exemplified by Hasuwa used a U6 promoter, but Hasuwa (page 230) also suggested using the tetracycline repressible U6 promoter of Ohkawa. Consequently, the technical feature of group I also is not a special technical feature, and as a result the construct of claim 23, which requires the technical features of group I and the special technical feature of either group II or group III, does not share a special technical feature with the construct of group I. Group IV is directed to a completely different product than that of groups I-III and does not share a technical feature (or special technical feature) with them.

Continuation of B. FIELDS SEARCHED Item 3:

MEDLINE, EMBASE, BIOSIS, CAPLUS, SCISEARCH, USPT, PGPB, DWPI, GENESEQ, PIR, SWISSPROT, SPTREMBL

search terms: RNAi, siRNA, shRNA, (interfering or interference) near RNA, (small or short) hairpin RNA, Pol III, U6TO, U6, tet operator, lac operator, tetO, tetracycline, lacO, loxP, SEQ ID NO: 53